Mobile Netw Appl
DOI 10.1007/s11036-015-0656-6

Energy Saving Scheme for Multicarrier HSPA+ Under Realistic Traffic Fluctuation

Maliha U. Jada[1] · Mario García-Lozano[2] · Jyri Hämäläinen[1]

Abstract In the near future, an increase in cellular network density is expected to be one of the main enablers to boost the system capacity. This development will lead to an increase in the network energy consumption. In this context, we propose an energy efficient dynamic scheme for HSDPA+ (High Speed Downlink Packet Access-Advanced) systems aggregating several carriers and which adapts dynamically to the network traffic. The scheme evaluates whether node-B deactivation is feasible without compromising the user flow throughput. Furthermore, instead of progressive de-activation of carriers and/or node-B switch-off, we evaluate the approach where feasible combination of inter-site distance and number of carriers is searched to obtain best savings. This is done by also considering the effect of transition delays between network configuration changes. The solution exploits the fact that re-activation of carriers might permit turning off other BSs earlier at relatively higher load than existing policies. Remote electrical downtilt is also considered as a means to maximize the utilization of higher modulation and coding schemes in the extended cells. This approach promises significant energy savings when compared with existing policies - not only for low traffic hours but also for medium load scenarios.

Keywords Multicarrier HSPA+ · Energy saving · Cell switch off · Carrier management

1 Introduction

Due to the increase in demand for mobile broadband services, mobile network vendors are preparing to challenging mobile data traffic increases. Whereas the initial forecasts indicated a 1000x traffic explosion between 2010 and 2020 [1], recent reports indicate predictions of a 57 % compound annual growth rate (CAGR) from 2014 to 2019 [2]. This is basically nothing new: According to [3] wireless capacity has already increased more than $10^6\times$ since 1957. Whereas $5\times$ comes from improvements in modulation and coding schemes (MCS), $1600\times$ increase is due to the reduction in cell sizes. It is widely accepted that the new capacity objective comes hand in hand with a further reduction in distances between transmission points. Network densification allows higher spatial reuse and so it allows higher area spectral efficiency measured in [bits/s/Hz/km^2]. On the other hand, considering that base stations (BSs) contribute the most to the energy consumption [4–6], future hyper-dense network deployments may negatively impact on the operational costs and carbon emissions.

It has become an important goal for industry and academia to reduce the energy consumption of mobile networks over coming years. Important effort has already been taken to address this issue. 3GPP launched several initiatives for energy saving in the radio access part of LTE [7, 8]. Also many research projects have put their focus on energy efficiency of mobile broadband systems, outstanding examples

✉ Mario García-Lozano
mariogarcia@tsc.upc.edu

Maliha U. Jada
maliha.jada@aalto.fi

Jyri Hämäläinen
jyri.hamalainen@aalto.fi

[1] Department of Communications and Networking, Aalto University, Espoo, Finland

[2] Department of Signal Theory and Communications, BarcelonaTech (UPC), Barcelona, Spain

are TREND [9], EARTH [10] and C2POWER [11]. Energy efficiency is also one of the key challenges in the evolution towards beyond fourth generation (4G) mobile communication systems. Yet, focusing in future systems is not enough since High Speed Packet Access (HSPA) and Long Term Evolution (LTE) will serve and coexist in the next decade, with probably a more tight integration in future releases of the standards [12]. In particular, HSPA is currently deployed in over 500 networks and it is expected to cover 90 % of the world's population by 2019 [13]. So it will serve the majority of subscribers during this decade while LTE continues its expansion in parallel and gain constantly large share of users.

Among the advantages in the latest releases of HSPA (HSPA+), multicarrier utilization is considered as an important performance booster [14] but it has not been extensively studied from the energy efficiency perspective so far. Given this, the focus of our study is in the reduction of energy consumption through dynamic usage of multiple carriers combined with the BS (node-B) switch-off.

Various BSs turn off strategies have been extensively studied as means for energy saving. Since cellular networks are dimensioned to correctly serve the traffic at the busy hour, the idea behind these strategies is to manage the activity of BSs in an energy-efficient manner while simultaneously being able to respond the traffic needs dynamically. Thus, the focus is on strategies where underutilized BSs are switched off during low traffic periods [6, 15–18]. This cell switch-off technique was also utilized in [19] where the focus was on the WCDMA energy savings through cell breathing while in [4] the same approach was considered along with the impact of femtocells on energy savings. The same technique was studied in [20] and the daily variation in traffic load was also modeled.

In order to guarantee coverage, switch-off is usually combined with a certain power increase in the remaining cells, but still providing a net gain in the global energy saving. However, this is not a straight-forward solution from a practical perspective: common control channels also require a power increase and electromagnetic exposure limits must be fulfilled [21]. Remote electrical downtilt lacks these problems. It positively impacts received powers and the coverage for common control channels could be expanded without increasing their power. A reduction of downtilt with respect to the horizontal can be used to increase the coverage range. Levels of interference are also modified, so the final geometry of the cell will change. For a complete analysis on the impact of downtilt over UMTS based systems, the reader is referred to [22]. More recently, BS cooperation has also been proposed to cover the newly introduced coverage holes when switch-off is applied [23].

Algorithms that minimize the energy consumption do also have an impact on the system capacity. The work in [24] studies these conflicting objectives and investigates cell switching off as a multiobjective optimization problem. This tradeoff should be carefully addressed, otherwise the applicability of a particular mechanism would be questionable. Yet, not many works consider the capacity issue in detail and many of the contributions just introduce a minimum signal-to-noise plus interference ratio (SINR) threshold, which allows to compute a minimum throughput or outage probability to be guaranteed. Consequently, capacity does not remain constant before and after the switch off. Indeed [25] strongly questions the applicability of cell switch off combined with power increases as a feasible solution for many scenarios.

Very few works evaluate energy saving gains obtained by advantageous use of the multi-carrier option. The works [26] and [27] respectively deal with HSPA and LTE when two carriers are aggregated and evaluate whether the additional carrier can be de-activated when load decreases and BSs are not powered off. LTE also offers a variety of channel bandwidth usage per carrier that ranges from 1.4 MHz to 20 MHz. Thus, several works have leveraged this feature to propose energy savings by using a dynamic adaptation of the bandwidth per carrier [28]. 3GPP initiated Rel12 efforts [29] to define a smaller bandwidth version of UMTS, so similar energy saving techniques could also be a matter of study in the context of HSPA.

The current work deals with the reduction of energy consumption in HSPA+ by means of a strategy that combines partial and complete node-B switch off with antenna downtilt. Utilization of multiple carriers is evaluated as an additional degree of freedom that allows more energy efficient network layouts. The number of available carriers is dynamically managed in combination with full BS turn off. This last action provides the highest energy saving. For this reason, instead of progressive de-activation of carriers until the eventual node-B turn off, we evaluate the combination (inter-site distance, number of carriers) that gives best energy saving. The solution exploits the fact that *activation* of previously shut-off carriers might permit turning off the BSs earlier at relatively higher load than existing policies. The new scheme promises significant energy savings when compared with existing policies, not only for low traffic hours but also for medium load scenarios. The work also provides an insight on the impact of transition times and delays to apply network updates, considering realistic load fluctuations.

The paper is organized in five sections. Section 2 discusses about the advantages and possibilities of multicarrier HSPA+. Section 3 describes the system model. In Section 4 we discuss about the BS shut-off scheme and Section 5 is devoted to results and discussion. Section 5 shows the

Table 1 Evolution of multicarrier HSDPA

Release	Name	Aggregation type
R8	Dual cell HSDPA	2 adjacent downlink carriers
R9	Dual band HSDPA	2 carriers from 2 different bands
R10	4C-HSDPA	Up to 4 carriers from one or 2 bands
R11	8C-HSDPA	Up to 8 carriers from one or 2 bands

effect of switching ON-OFF time delay on system performance and energy savings. Section 6 is devoted to results and discussions. Conclusions are drawn in Section 7.

2 Multicarrier HSPA+

Latest releases of HSPA offer numerous upgrade options with features such as higher order modulation, multi-carrier operation and multiple input multiple output (MIMO). Evolution from initial releases is smooth since MCS update and multicarrier are unexpensive features [30]. These advantages have motivated 65 % of HSPA operators to deploy HSPA+, as recorded December 2013 [13].

HSPA has evolved from a single carrier system to up to 8-carrier aggregation (8C-HSDPA). So, multicarrier operation can be supported in a variety of scenarios depending on the release, indicated in Table 1 for the downlink (HSDPA). Note that the uplink just allows dual cell since release 9. Multicarrier capability is an important advantage that affects the system performance [14, 31], in particular:

- It scales the user throughput with the number of carriers reaching a top theoretical speed of 672 Mbps on the downlink when combining 8C-HSDPA with 4×4 MIMO.
- It also improves spectrum utilization and the system capacity because of the load balancing between carriers.
- Multicarrier operation improves the user throughput for a given load at any location in the cell, even at the cell edge, where channel conditions are not good. Note that other techniques such as high order modulation combined with high rate coding or the transmission of parallel streams with MIMO require high SINRs. Furthermore, it is well known that every order of MIMO just doubles the rate only for users with good channel strength and no line of sight, while on cell edge MIMO just provides diversity or beamforming gain.

Regarding the availability of bandwidth, dual-carrier is currently mainstream solution. 8C-HSDPA is a likely option for scenarios in which bands from second generation (2G) systems are intensively refarmed or the use of unpaired bands as supplemental downlink is introduced [32]. On the other hand, scalable bandwidth for HSPA would also allow a more gradual refarming process and availability of new bandwidth pieces for aggregation [29]. However, the most interesting option would be a holistic management of the operator's spectrum blocks, with concurrent operation of GSM, HSPA and LTE that would allow an efficient resource sharing among technologies [33]. This multiaccess management can consider both quality of service (QoS) and energy efficiency as described herein.

3 System model

The BS shut-off scheme presents a well-defined solution for a problem of underutilized network elements. However, as previously stated, this action should be performed without compromising the system performance. This section presents the model to assess coverage and capacity dimensioning.

3.1 Coverage model

Let us assume the downlink of an HSPA+ system. At the link level, 30 modulation and coding schemes (MCS) are adaptively assigned by the scheduler based on the channel quality indicator (CQI) reported by user equipments (UEs). Given the channel condition and the available power for the high speed physical downlink shared channel (HS-PDSCH) $P_{\text{HS-PDSCH}}$, the scheduler selects the MCS that would guarantee a 10 % block error rate (BLER) for each user per transmission time interval (TTI).

The CQI reported on the uplink can be approximated using the SINR (γ in dB) at the UE for the required BLER as [34]:

$$CQI = \begin{cases} 0 & \text{if } \gamma \leq -16\text{dB} \\ \left\lfloor \frac{\gamma}{1.02} + 16.62 \right\rfloor & \text{if } -16\text{dB} < \gamma < 14\text{dB} \\ 30 & \text{if } 14\text{dB} \leq \gamma \end{cases} \quad (1)$$

Throughput of UE i depends on the number of allocated carriers and the SINR (γ_i in linear units) at each carrier f, given by:

$$\gamma_i = \frac{\frac{N_{\text{code}} P_{\text{code}}}{L_{s,i}}}{(1-\alpha)\frac{P_{\text{tot}} - P_{\text{code}}}{L_{s,i}} + \sum_{j \neq s} \left(\bar{\rho}_j \frac{P_{\text{tot}}}{L_{j,i}} \right) + P_{\text{N}}} SF_{16}, \quad (2)$$

where:

- For the sake of clarity, index referring to carrier f has been omitted.
- $L_{j,i}$ is the net loss in the link budget between cell j and UE i for carrier f. Note that index s refers to the serving cell.
- P_{tot} is the carrier transmission power. Without loss of generality, it is assumed equal in all cells of the scenario.

- Intercell interference is scaled by neighbouring cell load $\bar{\rho}$ at f (carrier activity factor).
- P_N is the noise power.
- P_{code} is the power allocated per HS-PDSCH code. Note that all codes intended for a certain UE shall be transmitted with equal power [35]. So, considering an allocation of N_{code} codes and a power P_{CCH} for the control channels that are present in f, then $P_{code} = \frac{P_{tot} - P_{CCH}}{N_{code}}$.
- SF_{16} is the HS-PDSCH (high-speed physical downlink shared channel) spreading factor of 16.
- The orthogonality factor α models the percentage of interference from other codes in the same orthogonal variable spreading factor (OVSF) tree. Our model assumes classic Rake receivers, in case of advanced devices (Type 2 and Type 3/3i) [36], their ability to partially suppress self-interference and interference from other users would be modelled by properly scaling the interfering power [37].

At the radio planning phase, a cell edge throughput is chosen and the link budget is adjusted so that the corresponding SINR (CQI) is guaranteed with a certain target probability p_t. Given that both useful and interfering average powers are log-normally distributed, the total interference is computed following the method in [38] for the summation of log-normal distributions. Coverage can be computed for any CQI and so, the boundary in which MCS k would be used with probability p_t can be estimated. This allows finding the area A_k in which k is allocated with probability $\geq p_t$. Figure 1 shows an example for a tri-sectorial layout with node-Bs regularly distributed using different inter-site distances (ISDs).

It is important to note that the node-B does not change its total power with each downtilt update. The shape of the final CQI rings largely depends on the antenna pattern, interference from neighbouring cells and downtilt, whose optimum value depends on the ISD. The example considers a multiband commercial antenna and downtilt angles are chosen following this criteria:

1. The minimum MCS is guaranteed at the cell edge.
2. After the previous constraint is met, the use of the highest MCS is maximized in the cell area.

Under these assumptions, the optimum angles for ISDs 750, 500 and 250 are 12.4°, 12.9° and 18.5° respectively.

Rings distribution will expand or reduce following the load in other cells. Figure 2 shows the pdf for CQIs 10 to 30 for 2, 4 or 8 carriers and the same cell load, and so different load per carrier $\bar{\rho}(f)$. This figure must be read jointly with Fig. 1 since the first has no numbers on the size of the different areas and the second does not represent how these values are geographically distributed Interference is spread among the different carriers and so the probability

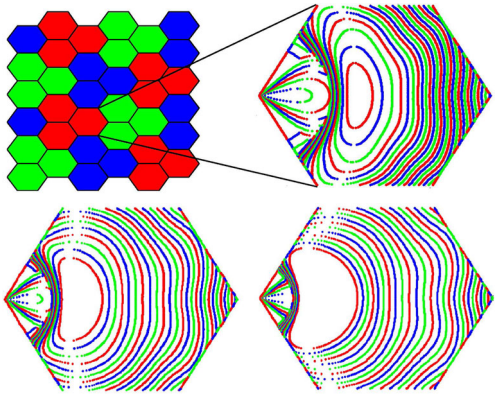


Fig. 1 Probabilistic CQI ring distribution in tri-sectorial regular layout. With ISD =250m *(top right)* ISD=500m *(bottom left)* ISD=750m *(bottom right)* at same cell load of 0.5 in each case and for their respective optimum downtilt angles

of allocating higher CQIs increases with the number of carriers. This has an impact on cell capacity and so, the next subsection is devoted to describe its model.

3.2 Capacity model

The capacity model largely follows [39]. We define cell capacity as the maximum traffic intensity that can be served by the cell without becoming saturated. Note that the cell load is evenly distributed among all carriers, so for the sake of clarity and without loss of generality, we will proceed the explanation assuming one single carrier and the index f will be omitted. It is important to note that a round robin scheduler is assumed. Therefore scheduling time is fairly shared among the users in the cell. Serving time depends on the cell load and allocated MCS, and so the download time is different for each user, more refined scheduling options would just shift absolute throughput values.

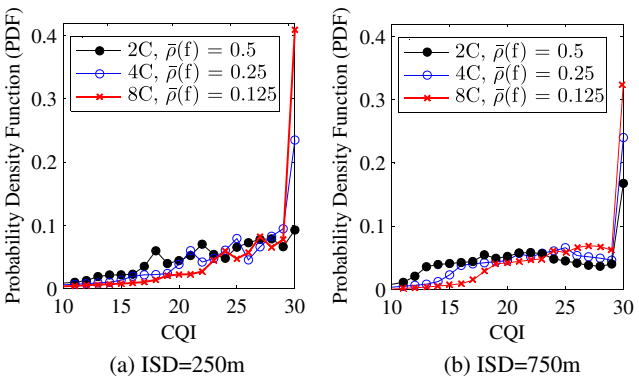


Fig. 2 CQI pdf for 2, 4 and 8 carriers and same cell load

3.2.1 Cell capacity

Let's assume the traffic to be uniformly distributed in the cell. Data flows arrive according to a Poisson process with rate λ per area unit. Flow sizes are independent and identically distributed with average size $\mathrm{E}(\sigma)$. So, the cell load or fraction of time in which the scheduler must be active is:

$$\bar{\rho} = \lambda A_{\text{cell}} \times \sum_{k=1}^{30} \frac{\mathrm{E}(\sigma)}{c_k} p_k \leq 1 \,, \tag{3}$$

where A_{cell} is the cell area, c_k is the achievable date rate associated to MCS k, and p_k is the probability of using MCS k, $p_k = A_k/A_{\text{cell}}$, being A_k the area in which MCS k is allocated. Note that A_k corresponds with the ring in which CQI k was reported by mobile users and so, it is just a fraction of the total A_{cell}. Since the cell load is bounded to one, the maximum throughput that can be served ($\bar{\rho} = 1$), or cell capacity is:

$$\bar{C} = \left(\sum_{k=1}^{30} \frac{p_k}{c_k} \right)^{-1} . \tag{4}$$

At any given load, the observed throughput (served) would be given by $\bar{\rho} \times \bar{C}$.

3.2.2 Flow throughput

Actions to provide energy savings should not compromise the QoS and so the user flow throughput should not be altered. Hence, this has been used as performance metric. Since the scheduler is shared among the users in the cell, serving time depends on the cell load and allocated MCS, and so the download time is different for all users. The contribution to the cell load (fraction of time that should be allocated by the scheduler) from users at A_k is given by

$$\bar{\rho}_k = \frac{\lambda A_k \times \mathrm{E}(\sigma)}{c_k} . \tag{5}$$

Hence, considering the definition of $\bar{\rho}$ in Eq. 3, it can be observed that it is immediate to also express it as $\bar{\rho} = \sum_k \bar{\rho}_k$. Given that all users in the cell share the same scheduler, by using Little law's the mean flow duration t_k for a user in A_k can be computed $t_k = N_k/\lambda A_k$ where N_k is the average number of users in A_k. Then the flow throughput τ_k for users being served with MCS k is:

$$\tau_k = \frac{\mathrm{E}(\sigma)}{t_k} = \frac{\lambda A_k \times \mathrm{E}(\sigma)}{N_k} \,, \tag{6}$$

Considering the underlaying Markov process [39], it can be found the stationary distribution of the number of active users in each A_k and its average value, $N_k = \frac{\bar{\rho}_k}{1-\bar{\rho}}$, which yields:

$$\tau_k = c_k(1 - \bar{\rho}) \,, \tag{7}$$

and the average flow throughput at cell level:

$$\bar{\tau} = \sum_{k=1}^{30} p_k c_k (1 - \bar{\rho}) \,, \tag{8}$$

where $\bar{\rho}$ captures the own cell load and p_k is affected by the load in neighbouring cells, which modifies SINR values, CQI rings and so A_k values $\forall k$.

4 Node-B shut off scheme and energy model

The network configuration is characterized by the duplet (ISD, N_{car}). The transition between two pairs of values, i.e. the transition of a network configuration (in terms of ISD and active number of carriers) takes place at certain load thresholds in which the network can be reconfigured to achieve a more energy efficient layout without compromising the required average flow throughput. For any load z the best network configuration is the one which satisfies two conditions:

1. The load z is such that with the new configuration the average flow throughput $\bar{\tau}$ is maintained: $\bar{\tau}_{\text{new}}(z) \geq \bar{\tau}_{\text{initial}}(z)$.
2. The new configuration at load z allows a lower energy consumed per unit area E/A (kWh/km^2) than the old configuration: $E/A_{\text{new}}(z) \leq E/A_{\text{old}}(z)$.

Thus that value of load becomes the threshold level at which the network transition should take place. It is important to note that the mechanism not only reduces active sites or carriers progressively. The main contribution is to combine these two mechanisms. For example, at some point it is more beneficial a combination of de-activation of sites + activation of carriers. This is because the power saving due to a complete switch off is dominant over the activation of a certain number of carriers. So, that action would be triggered if the load is low enough to increase the ISD but extra radio resources are yet needed to accommodate the new (higher) cell load. This provides more granularity in the set of possible optimization actions and so higher energy saving gains can be obtained. The set of all possible actions are indicated in the two first columns of Table 2, where the arrow indicates the initial and end of the transition. For example, ON $\rightarrow$ OFF in the first column means a transition in which base stations are deactivated. The rest of the columns will be explained during Section 5.

Any BS shut off is followed by an update of downtilt angles in remaining cells to guarantee coverage for cell edge users and to maximize the use of the highest possible MCS under the new conditions. Downtilt angle is reduced (with respect to the horizontal) with the increasing ISD. Note that

Table 2 Network configuration state that is chosen in the energy model for all possible network transition combinations

Base stations	Frequency carriers	N_s^{STATE}	N_{car}^{STATE}	T_{delay}
No-Change	No-Change			null
OFF → ON	No-Change	N_s^{New}	N_{car}^{New}	long
No-Change	OFF → ON			short
OFF → ON	OFF → ON			long
OFF → ON	ON → OFF	N_s^{New}	N_{car}^{Prev}	long
No-Change	ON → OFF			short
ON → OFF	OFF → ON	N_s^{Prev}	N_{car}^{New}	long
ON → OFF	No-Change			long
ON → OFF	ON → OFF	N_s^{Prev}	N_{car}^{Prev}	long

transmission power remains unchanged no matter the new ISD.

Note that whenever sites are switched off, a readjustment of the complete grid of ISD is performed. Thus the method is applied over a set of base stations in a particular area of the network. Common load variations along the day are followed for that particular area. Thus, very short term fluctuations that individually happen at the sector level are not an objective of this method.

After performing node-B shut off, the higher load of the new expanded cells is again evenly distributed over the total frequency resources. This high load includes the user traffic of the switched off node-Bs, which has to be accommodated by the remaining active ones. Considering υ as the ratio $ISD^{new}/ISD^{initial}$ and that the cell area A_{cell} is proportional to the cell square radius, from Eq. 3, the relation between the cell load with new ISD $\bar{\rho}^{new}$ and with initial ISD $\bar{\rho}^{initial}$ is given by,

$$\bar{\rho}^{new} = \upsilon^2 \times \bar{\rho}^{initial} . \tag{9}$$

Although the use of more carriers will account for a certain increase in energy consumption, the saving for switching off some BSs is much higher.

The metric used for the analysis of energy consumption is energy consumed per unit area (E/A). Assuming an entire parallel system at the node-B to handle each carrier, the energy consumed per unit area (kWh/km^2) is given as [4, 19, 26, 40]:

$$E/A = \frac{N_s \cdot N_{sec} \cdot N_{car} \cdot \left[P_{oper} + (\bar{\rho}(f, N_s, N_{car}) \cdot P_{in})\right] \cdot T}{A_{tot}} , \tag{10}$$

where:

- N_s, N_{sec}, N_{car} are the number of sites, sectors (or cell) per site and carriers per sector (or cell) respectively.
- P_{oper} is the operational power, which is the load independent power to operate the node-B and includes baseband processing.
- $\bar{\rho}(f, N_s, N_{car})$ is the cell load that varies from 0.1 (relevant to the minimum transmit pilot power) to 1.0 referring to the cell at full load. Note the dependence among variables, any change in N_s and/or N_{car} implies the corresponding update in $\bar{\rho}$.
- P_{in} is the power consumed to eventually obtain the required power at the antenna connector (P_{Tx}). It scales linearly with the load $\bar{\rho}$ but it is not P_{Tx}. P_{in} is based on the rated power of the power amplifier (PA); such that $P_{in} = P_{Tx}/PA_{Eff}$, where PA_{Eff} is the power amplifier efficiency, which we adapt from [40] as 35 %.
- T is the time duration in which the particular load remains in the piece of network under study.
- A_{tot} is the total evaluated area containing N_s sites.

For the case of UMTS, power consumption of currently deployed base stations is mainly dominated by the operational (static) part, the amount of dynamic power is negligible, around 3 % [41]. In our specific calculations we adapt from [19] the UMTS macro base station specific values $P_{oper} = 157$W, $P_{in} = 57$W in this last case, conveniently weighted by the current load. This means a less unbalanced relationship P_{oper}/P_{in} with respect to currently widespread technology. It is important to note that this relationship will directly affect the achievable power savings, more details are provided in the results section.

5 Impact of switching on-off time-delay on QoS and energy savings

The transition in the network configuration, understood as the number of active BSs and frequency carriers, does not occur instantly. There is always some time delay involved in switching ON-OFF the BSs and carriers. During these time periods, when the network is considered to be in a transitional state, the users might have to experience a non-optimal QoS and, on the operational side, the service provider might have to bear some extra energy expenses [42, 43]. These switching ON-OFF time periods could be of few seconds to few minutes depending upon the hardware and system capabilities [44].

Considering the importance of this issue, usually missed in the existent literature, this work will also quantify the influence of such time-delays on the QoS and energy consumption of the network.

Whenever the load levels justify a certain BS or carrier being from OFF to ON state, the throughput during the transition is given by the previous network configuration and the new load. This is because the new resources

are not fully utilizable until the end of their activation. This means that in some cases the network configuration will operate under non-optimal conditions with the corresponding throughput loss. On the other hand, if changes are very much anticipated, the system would be implementing less energy efficient configurations. In a similar manner, when deactivations are required, the BS or frequency resources become unavailable instantly after triggering the OFF decision but energy savings do not occur until the equipment is fully switched off and so until the end of the transition.

Given this, the energy consumption model requires to be updated. In particular, the energy per unit area for the transitional time period is given as:

$$(E/A)^{\text{trans}} = \frac{1}{A_{\text{tot}}} \cdot N_{\text{sec}} \cdot N_{\text{s}}^{\text{STATE}} \cdot N_{\text{car}}^{\text{STATE}} \cdot \left[P_{\text{oper}} + \left(\bar{\rho} \left(f, N_{\text{s}}^{\text{STATE}}, N_{\text{car}}^{\text{STATE}} \right) \cdot P_{\text{in}} \right) \right] \cdot T_{\text{delay}}, \quad (11)$$

where, $N_{\text{s}}^{\text{STATE}}$ and $N_{\text{car}}^{\text{STATE}}$ are the number of active sites and carriers per sector for a particular configuration state which is decided based on the action taken: turn ON or OFF of BS sites and/or carriers. Thus, if there is no change or new base stations are switched on, $N_{\text{s}}^{\text{STATE}}$ equals the new number of active sites, since operational power consumption increases since the very first moment of activation, no matter the time it takes for the node-B to be completely operative to start serving users. Transition time would be shorter when there is no node-B (de-)activation, since swithing on and off a transceiver is far faster and simpler than a complete station. Table 2 summarizes the network configuration state (current or previous) that is chosen in the model for all possible network transition combinations (switching ON-OFF BSs & carriers) and where $N_{\text{x}}^{\text{New/Prev}}$ indicates if the new or previous number of resources 'x' appear in the transition energy model.

Regarding, the power provided at the antenna connector P_{in}, it must be noted that the load value $\bar{\rho}$ is also updated conveniently. For example, if an activation occurs, the load cannot be shared among the new increased number of elements (sites and or carriers) until the end of the transition time. On the other hand, if de-activation occurs, the resources are not available any more from the very beginning, no matter how long it takes to de-activate the element.

Given this, it is straight forward that the total energy consumed for the current load including the transition period is given as,

$$(E/A)' = (E/A) \cdot \frac{T - T_{\text{delay}}}{T} + (E/A)^{\text{trans}}. \quad (12)$$

Table 3 Evaluated scenarios

	Scenario 1	Scenario 2	Scenario 3
Targeted $\bar{\tau}$	5.76 Mbps	21.80 Mbps	60.53 Mbps

6 Results

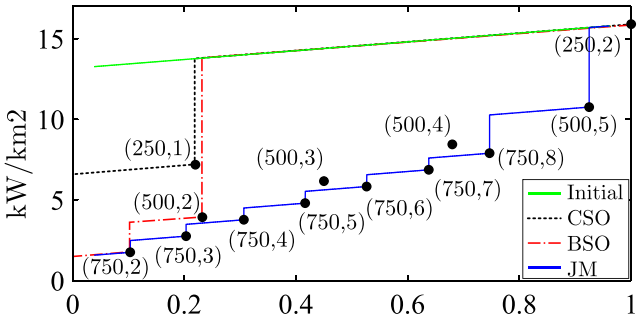
In order to quantify the gains that can be achieved by an intelligent joint management of carriers and node-Bs, the system performance is evaluated in terms of average flow throughput ($\bar{\tau}$). Three cases have been evaluated: 2, 4 and 8 carriers are initially used to serve an aggregated cell load of 1. This load is evenly distributed among the carriers, $\bar{\rho} = 0.5$, 0.25, and 0.125. Given this, three scenarios are defined considering the $\bar{\tau}$ value to be respected (Table 3). Note that as the load is distributed among more carriers, the average cell flow throughput increases due to a reduction in the utilization of low MCSs and of course to the availability of more radio resources.

The reference network has node-Bs regularly deployed and considering an ISD=250 m. Therefore after a first shut off, the new ISD would be 500 m, and a second implies ISD=750 m. Other network parameters are provided in Table 4.

Figure 3 represents the power consumption per unit area for decreasing cell load values and showing the transition points that should be used to guarantee the target user flow throughput after the network update. In each subplot four cases are represented:

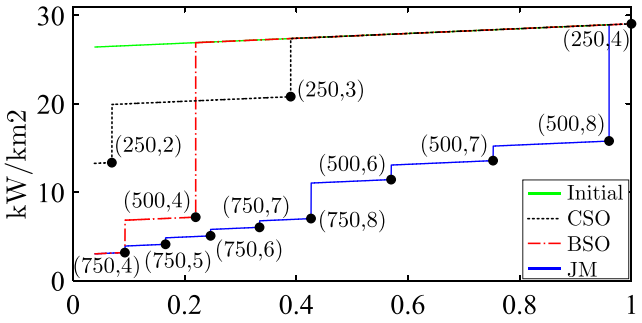
Table 4 Network parameters

Parameter	Value
Operating bands	2100 MHz, 900 MHz
Carrier bandwidth	5 MHz
Inter-Site distances	250 m, 500 m, 750 m
Number of sites, for each ISD	108, 27, 12
BS transmission power/sector/carrier	20 W
Transmission power per user	17 W
Control overhead	15 %
BS antenna gain	18 dB
Body loss	2 dB
Cable and connection loss	4 dB
Noise power	−100.13 dBm
Propagation model	Okumura-Hata
Shadow fading std. deviation	8 dB
Cell edge coverage probability	0.99
$T_{\text{delay}} = \text{long}$	2 minutes
$T_{\text{delay}} = \text{short}$	1 minute



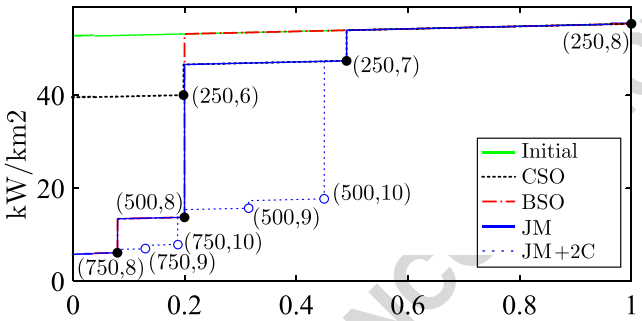
Cell load of initial network with ISD=250 m

(a) Scenario 1



Cell load of initial network with ISD=250 m

(b) Scenario 2



Cell load of initial network with ISD=250 m

(c) Scenario 3

Fig. 3 Power consumption per unit area for different combinations of the duplet (ISD, number of carriers) and for different load values. The different curves follow the duplets that would imply the lowest power consumption for each strategy. Example from fig **(a)**: When the cell load is one, (250, 2) is required but a transition to (500, 5) can be executed after a slight decrease in the load

- **Initial**: Power consumption under the initial network configuration, without changes. It can be seen that it just depends on the system load and so the power consumption is just slightly reduced.
- **BSO**: Base station shut off. Classic model in which node-Bs are successively shut off whenever the load allows to still keep the target flow throughput. Thus, it is important to note that in this case there is no carrier management. The only action that can be taken is a progressive deactivation of nodes to increase the ISD.
- **CSO**: Carrier shut off. Generalization of the DC-HSDPA case in [26] for a multi-carrier case, carriers are progressively shut off with load reduction. Unlike the previous case, there are no actions taken over the node-Bs, which always remain active. So the best case in terms of energy consumption that can be achieved is the duplet (250, 1) meaning ISD of 250 m and only 1 active carrier per sector.
- **JM**: Joint management. The proposal of the current work. Power off of BSs and carriers are jointly managed and re-activation of carriers is a valid option if that justifies earlier full BSs shut off and so a net energy saving.

Each tag in the plot shows the transition points in terms of (ISD, number of active carriers). Since the load is progressively reduced, the pictures should be read from right to left. For example, for the BSO case in Scenario 1, the transition points evolve as $(250, 2) \rightarrow (500, 2) \rightarrow (750, 2)$, note how the last case can only be implemented for cell loads of 10 %, meaning a 5 % of load per carrier.

It is important to note that the absolute values of power consumption and load triggering a network change closely depend on the static and dynamic required powers, P_{oper} and P_{in} in Eq. 10. If we assume a different power model, in which some static parts are also load dependent, then P_{oper} would be lower and savings in each transition would be shorter. On the other hand, if the dynamic part is more sensitive to load changes, the slope of each segment in the plot would be higher, thus leading to a faster decrease in power consumption.

The joint management allows earlier BS shut off and transition points fall below the other options, thus having clearly less power consumption without performance degradation. Note that for classic methods, transitions always happen for low load values, which shows that at medium load levels it is not possible to just shut off BSs without user throughput impairment. It can be seen how JM allows using ISD=750 m as soon as the cell load falls below 0.8. For Scenario 2, the ISD can be increased from 250 to 500 for high loads, and 750 m can be used once the load falls below 0.5. Scenario 3 is the most restrictive since it starts with the maximum possible carriers at the current HSPA+ standard. So there is less flexibility with respect to the other cases and the savings are just slightly better. For illustrative purposes, it has been included the off-standard case in which up to 10 carriers are used, it can be seen how energy savings are again important. This way, multiaccess energy saving

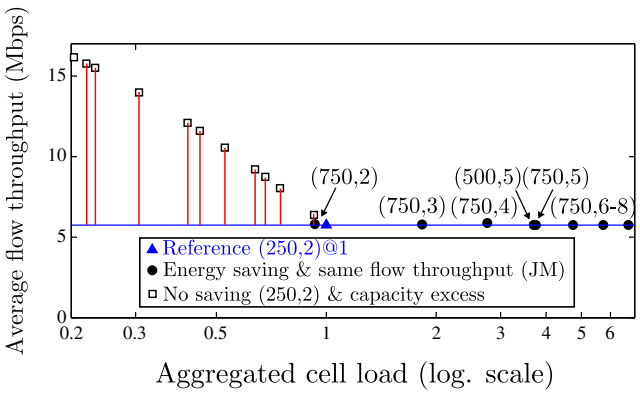


Fig. 4 Transition of cell configuration from initial network setup (scenario 1) to new setups at specific load values and maintaining the QoS requirements (5.75 Mbps)

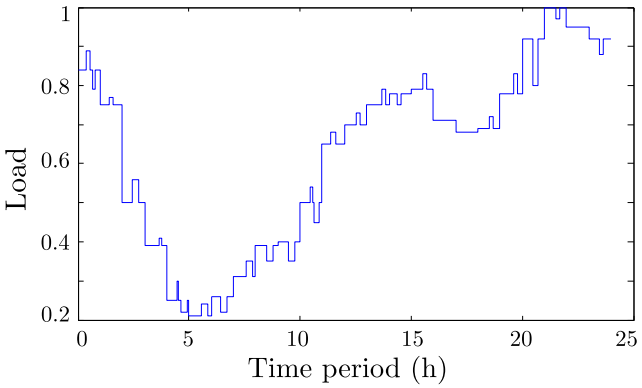


Fig. 5 Traffic load fluctuations

mechanisms that manage the pool of resources among several systems would make the most of each system load variations.

It is important to note that the horizontal axis represents the equivalent cell load that would be obtained if the network remained unchanged. But obviously, after carrier and/or node-B switch off, the cell load changes. For example, initially the load is 1 (0.5 per carrier) and it is not until it is reduced to 0.92 that important energy savings are possible, so we transition from (250, 2)@0.92 to (500, 5)@3.7. Please note that the value after @ represents the cell load when varying the number of carriers and the ISD distance. Recall that since the load per carrier is bounded to 1, the final aggregated cell load value can be > 1. Besides, it is clear that the cell load increases due to its expansion and the new users to be served, but the QoS is respected, since both (250, 2)@1 and (500, 5)@3.7 provide the same flow throughput.

In order to illustrate how load evolves with every change, Fig. 4 represents the average flow throughput as a function of the aggregated cell load for each configuration proposed by JM (solid symbols). Note the logarithmic scale in the horizontal axis to improve readability. Their evolution (Fig. 3a) is as follows: (250, 2)@1 $\rightarrow$ (500, 5)@3.7 $\rightarrow$ (750, 8)@6.72 and so on. If no energy savings mechanisms are implemented, in other words, if we remain with the dense node-B deployment, an excess in capacity would be obtained due to load decrement. These situations are represented by empty symbols.

Given the previous results, in the following we consider a realistic profile of daily HSDPA traffic (load) [26] (Fig. 5) and evaluate energy consumption and corresponding savings along time.

Figure 6 represents results for scenarios 1 and 2. In case of Scenario 1, the total energy saving percentage is 45.4 % with JM, whereas it is just 2.8 % with BSO and 1.8% with CSO. For Scenario 2, gains increase up to 55.8 % for JM, and 2.9 %, 5.9 % for BSO and CSO respectively. Scenario 3 had an equal saving of just 3.5 % in CSO and JM, with no possible gain with BSO. As previously mentioned this is because scenario 3 is very restrictive and requires a flow throughput of 60.53 Mbps. In the hypothetical off-standard case with up to 10 available carriers, energy savings with JM would reach 19.9 %. From Fig. 6 it is also noticeable how small reductions in the load can lead to important savings as it happens with cell load values around 60 %. So we can conclude that even at mid-high values, interesting savings are possible when applying the JM approach.

Given the T_{delay} that takes to switch on/off a node-B completely, it is clear that these type of strategies cannot follow the short term fluctuations in the load demands. Besides, as it was previously explained, this delay will also imply a non-optimal operation of the network during transition

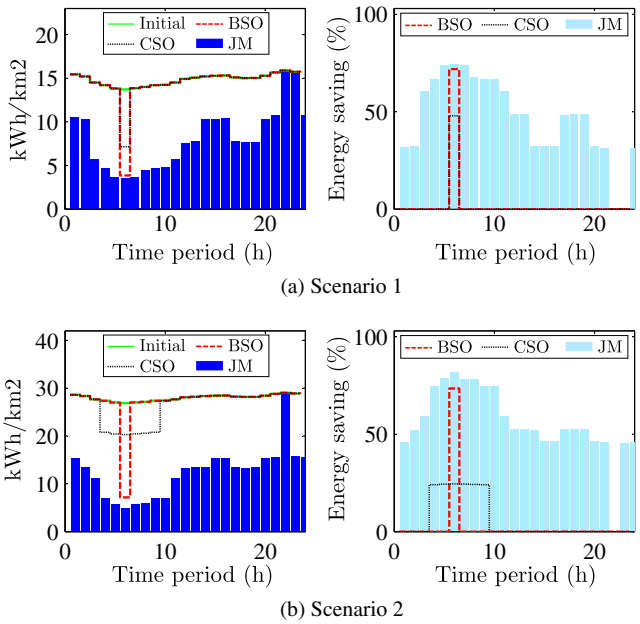


Fig. 6 Comparison between energy savings (%) of BSO, CSO and JM

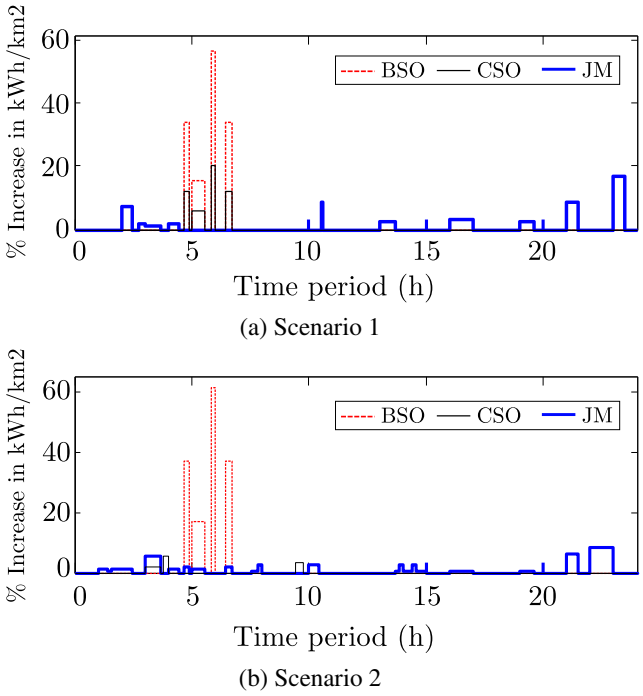


Fig. 7 Percentage increase in energy consumption

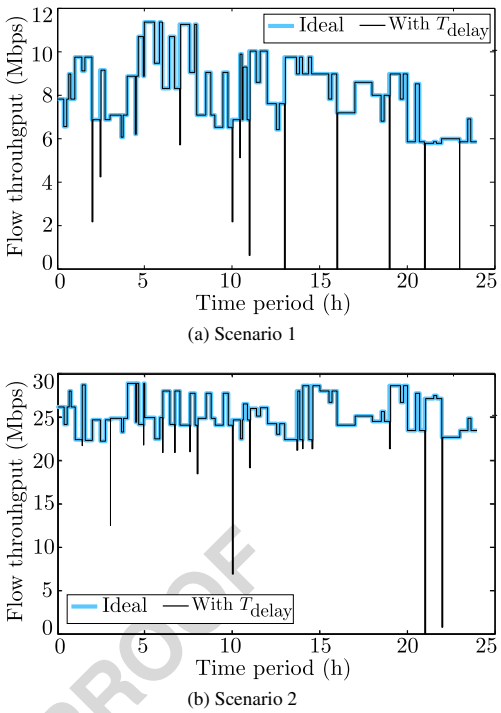


Fig. 8 Flow throughput evolution

times, in some cases in terms of QoS, in others in terms of energy efficiency. The last set of results aim at quantifying the energy and throughput variations due to the fast load fluctuations.

Figure 7 indicates the percentage increase in the energy consumption at some of the transitions, for each scheme (BSO, CSO, JM) for the first two scenarios. We note that for scenario 1 with JM scheme the percentage increase in overall energy consumption is 1.2 % in comparison to the case when delay was not considered. Similarly for BSO and CSO the percentage increase in energy consumption is 0.4 % and 0.25 % respectively. For scenario 2 with JM, BSO and CSO schemes the percentage increase in energy consumption is 1 %, 0.4 % and 0.1 % respectively and they are negligible for scenario 3 due to the low number of transitions. The proposed scheme performs more changes, adapts in a finer manner to load changes. But as it is shown, after considering the effect of delay transitions and short term load fluctuations, the gains are slightly reduced. Nevertheless, the impact is far lower than the resulting gains.

To end this section, Fig. 8 shows the degradation in flow throughput values during the transition time. Again, although this is not a very significant time period, from the user perspective the degradation in QoS for such small time periods cannot be ignored. The results are plotted for the new proposal JM since it is the case with more transitions and so, the strategy that is more affected. In particular, the plot compares the throughput evolution without considering time delays (ideal) and including them. It can be observed, how sub-optimum operation leads to important but brief reductions, in some cases reaching zero. This case indicates those situations in which the load would be higher than one, meaning that more than the available frequency resources are needed to have an stable system and be able to serve the offered load. The cases BSO and CSO are less affected since they also imply less network changes. Nevertheless, the energy saving is still much better in JM than existent proposals, as previously discussed.

7 Conclusions

In this paper we investigated the potential energy savings by shutting off the BSs through the dynamic use of multiple carriers in HSDPA. We have proposed an energy saving scheme in which fewer or additional carriers have been used depending upon the network traffic variations. This is combined with remote electrical downtilts to partially cope with the use of a higher number of lower MCSs. Instead of just guaranteeing a power threshold at the cell edge, or an outage probability threshold for data traffic, it is more interesting to ensure that QoS remains unchanged whenever a node-B and/or carrier is shut-off, for this reason the study considers user flow throughput as the performance metric

to be respected, which is closely affected by load variations due to cell expansions. Comparison to schemes that progressively shut off network elements (BSO and CSO) has been done, showing clear energy savings with the JM approach. The study includes the effects of transition times and delays required to switch between network configurations. Since JM is a strategy with more frequent updates, the negative effects of such delays in terms of QoS and energy savings are more present but still far from counteracting the gains.

The main challenge to make the adaptation efficient and flexible is that load fluctuations should be correctly followed. Reiterative traffic patterns can be assessed along time but abnormal temporal or spatial variations could be included in the system by means of a pattern recognition system, e.g. a fuzzy logic based system or a neural network. Further efforts are required in this direction.

Acknowledgments This work was supported in part by Academy of Finland under grant 284634. The work by Mario García-Lozano is funded by the Spanish National Science Council through project TEC2014-60258-C2-2-R.

References

1. NSN (2011) 2020: The ubiquitous heterogeneous network - Beyond 4G, tech. rep., ITU Kaleidoscope, NSN
2. CISCO VNI forecast (2015) Technical Report, CISCO
3. Chandrasekhar V, Andrews J, Gatherer A (2008) Femtocell networks: a survey. IEEE Comm Mag 46:59–67
4. Jada M, Hossain M, Hämäläinen J, Jäntti R (2010) Impact of femtocells to the WCDMA network energy efficiency. In: 3rd IEEE Broadband Network and Multimedia Technology. IC-BNMT, Beijing (China)
5. Jada M, Hossain M, Hämäläinen J, Jäntti R (2010) Power efficiency model for mobile access network. In: 21st IEEE Personal, Indoor and Mobile Radio Communications Workshops. PIMRC Workshops, Istanbul (Turkey)
6. Marsan M, Chiaraviglio L, Ciullo D, Meo M (2009) Optimal energy savings in cellular access networks. In: IEEE Int. Conf. on Comm. Workshops. ICC Workshops, Dresden (Germany)
7. 3GPP (2014) TR 36.927 (Release 12) - Evolved Universal Terrestrial Radio Access (E-UTRA); Potential Solutions for Energy Saving for E-UTRAN, Technical Report, 3GPP
8. 3GPP (2014) TR 36.887 (Release 12) - Study on Energy Saving Enhancement for E-UTRAN, Technical Report, 3GPP
9. FP7 EU project TREND, Towards Real Energy-efficient Network Design. http://www.fp7-trend.eu/. Accessed: 2015-07-01
10. FP7 EU project EARTH, Energy Aware Radio and Network Technologies. https://www.ict-earth.eu/. Accessed: 2015-07-01
11. FP7 EU project C2POWER, Cognitive Radio and Cooperative Strategies for Power Saving in Multi-standard Wireless Devices. http://www.ict-c2power.eu/. Accessed: 2015-07-01
12. Yang R, Chang Y, Xu W, Yang D (2013) Hybrid multi-radio transmission diversity scheme to improve wireless TCP Performance in an Integrated LTE and HSDPA Networks. In: IEEE Vehicular Tech. Conf. VTC Spring, Dresden (Germany)
13. 4G Americas (2014) White paper on 4G mobile broadband evolution: 3GPP Release 11 & Release 12 and Beyond, tech. rep., 4G Americas
14. Johansson K, Bergman J, Gerstenberger D, Blomgren M, Wallen A (2009) Multi-carrier HSPA evolution. In: IEEE Vehicular Tech. Conf. VTC Spring, Barcelona (Spain)
15. Gong J, Zhou S, Niu Z, Yang P (2010) Traffic-aware base station sleeping in dense cellular networks. In: Int. Workshop on Quality of Service. IWQoS, Beijing (China)
16. Niu Z (2011) TANGO: traffic-aware network planning and green operation. IEEE Wireless Comm 18:25–29
17. Oh E, Krishnamachari B (2010) Energy savings through dynamic base station switching in cellular wireless access networks. In: Proc. of IEEE Global Telecommunications Conference. GLOBECOM, Miami (United States)
18. Zhou S, Gong J, Yang Z, Niu Z, Yang P (2009) Green mobile access network with dynamic base station energy saving. In: MobiCom, Beijing
19. Micallef G, Mogensen P, Scheck HO (2010) Cell size breathing and possibilities to introduce cell sleep mode. In: European Wireless Conference, (Italy)
20. Richter F, Fehske AJ, Fettweis GP (2009) Energy efficiency aspects of base station deployment strategies for cellular networks. In: IEEE Vehicular Technology Conference, (USA)
21. Chiaraviglio L, Ciullo D, Meo M, Marsan M (2009) Energy-efficient management of UMTS access networks. In: Int. Teletraffic Congress. ITC, Paris (France)
22. Garcia-Lozano M, Ruiz S (2004) Effects of downtilting on RRM parameters. In: IEEE Int. Symp. on Personal, Indoor and Mobile Radio Comm. PIMRC, Barcelona (Spain)
23. Han F, Safar Z, Lin W, Chen Y, Liu K (2012) Energy-efficient cellular network operation via base station cooperation. In: IEEE Int. Conf. on Communications. ICC, Ottawa (Canada)
24. González D, Yanikomeroglu GH, Garcia-Lozano M, Ruiz S (2014) A novel multiobjective framework for cell switch-off in dense cellular networks. In: IEEE Int. Conf. on Comm. ICC, Sydney (Australia)
25. Wang X, Krishnamurthy P, Tipper D (2012) Cell sleeping for energy efficiency in cellular networks: is it viable? In: IEEE Wireless Comm. and Networking Conf. WCNC, Paris (France)
26. Micallef G, Mogensen P, Scheck H-O (2010) Dual-cell HSDPA for network energy saving. In: IEEE Vehicular Tech. Conf. VTC Spring, Taipei (Taiwan)
27. Chung Y-L (2013) Novel energy-efficient transmissions in 4G downlink networks. In: Int. Conf. on Innovative Comp. Tech. INTECH, London (UK)
28. Ambrosy A, Wilhelm M, Wajda W, Blume O (2012) Dynamic bandwidth management for energy savings in wireless base stations. In: IEEE GLOBECOM
29. 3GPP (2014) TR 25.701 v12.1.0 (Release 12) - Study on scalable UMTS Frequency Division Duplex (FDD) Bandwidth, Technical Report, 3GPP
30. Borkowski J, Husikyan L, Husikyan H (2012) HSPA evolution with CAPEX considerations. In: Int. Symp. on Comm. Systems, Networks Digital Signal Processing. CSNDSP, Poznan (Poland)
31. Bonald T, Elayoubi SE, El Falou A, Landre JB (2011) Radio capacity improvement with HSPA+ dual-cell. In: IEEE Int. Conf. on Communications. ICC, Kyoto (Japan)
32. 3GPP (2014) RP-140092 - Revised Work Item: L-band for Supplemental Downlink in E-UTRA and UTRA, tech. rep., 3GPP
33. NSN (2014) Answering the network energy challenge (whitepaper), tech. rep., NSN

34. Brouwer F, de Bruin I, Silva J, Souto N, Cercas F, Correia A (2004) Usage of link-level performance indicators for HSDPA network-level simulations in E-UMTS. In: Int. Symp. on Spread Spectrum Techniques and Applications. ISSSTA, Sydney (Australia)
35. 3GPP (2014) TR 25.214 v11.8.0 (Release 11) - Physical layer procedures (FDD), Technical Specification, 3GPP
36. 3GPP (2014) TR 25.101 v12.3.0 (Release 12) - User Equipment (UE) Radio Transmission and Reception (FDD), Technical Report, 3GPP
37. Rupp M, Caban S, Mehlfuhrer C, Wrulich M (2011) Evaluation of HSDPA and LTE From Testbed Measurements to System Level Performance. Wiley
38. Beeke K (2007) Spectrum planning - analysis of methods for the summation of log-normal distributions, EBU Technical Review
39. Bonald T, Proutière A (2003) Wireless downlink data channels: user performance and cell dimensioning. In: Annual Int. Conf. on Mobile Comp. and Networking. MOBICOM, San Diego, CA (USA)
40. Arnold O, Richter F, Fettweis G, Blume O (2010) Power consumption modeling of different base station types in heterogeneous cellular networks. In: Future Network and Mobile Summit, Florence (Italy)
41. Corliano A, Hufschmid M (2008) Energieverbracuh der mobilen kommunikation - energy consumption in mobile communications, in German, Technical Report, Federal Office of Energy in Switzerland (Technical Report)
42. Saker L, Elayoubi SE (2010) Sleep mode implementation issues in green base stations. In: 21st Annual IEEE International Symposium on Personal. Indoor and Mobile Radio Communications, Turkey
43. Marsan MA, Chiaraviglio L, Ciullo D, Meo M (2011) Switch-off transients in cellular access networks with sleep modes. In: Proceedings of IEEE International Conference on Communications. ICC, Kyoto (Japan)
44. Ohmann D, Fehske A, Fettweis G (2013) Transient flow level models for interference-coupled cellular networks. In: Fifty-first Annual Allerton Conference, USA